Week 2 Outlines

- Introductory Statistics
    - General Knowledge and Terms
    - Continuity of Data
    - Frequency Distribution
    - First Order of Average Measurement (Central Tendency)
        - Mean
        - Median
        - Mode
    - Data Positioning
        - Quartile, Decile, Percentile, and n-tile
- Assignments

    แบบฝึกหัดใน RATH Center (ไม่ต้องส่ง)

    ทบทวนคณิตศาสตร์พีชคณิต (Algebra-Precalculus)

Week 3 Scopes

- Second Order of Average Measurement
    - Deviation
    - etc.

Introductory Statistics

ให้ใช้แบบฝึกหัดวิชาคณิตศาสตร์มัธยมศึกษาตอนปลายของ RATH Center เป็นเอกสารประกอบ
ใช้ฝึกทำโจทย์ต่าง ๆ มีประโยชน์ทั้งการทำงานจริงและการสอบวิชาคณิตศาสตร์

General Knowledge

1. True/False-Correct

_______ a. Phone number is quantitative data.

_______ b. District population data from district office are primary data.

_______ c. Data average is a way of data visualization.

_______ d. Visualization makes data understandable.

2. Which one(s) of these should not affect statistical decision?

a. Data b. News

c. Information d. Belief

3. Numerical data is mostly or usually identified as (quantitative / qualitative) data.

For 4. And 5.:

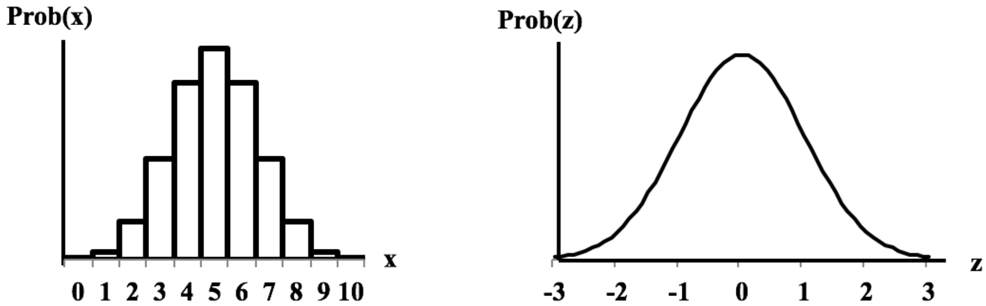
	Diameter	**Length**
Nozzle A	12.80	108
Nozzle B	37.95	46.7

4. The following data lack of ____ .
(May lack more than one)

5. Does it has consistency? _______

ความต่อเนื่องของข้อมูล (Continuity of Data)

โดยทั่วไป ข้อมูล (เจาะจงไปที่ข้อมูลเชิงปริมาณ เพราะการแบ่งประเภทของข้อมูลตามลักษณะมีความคลุมเครือสูง จึงไม่แบ่งชัดเจนเป็นเส้นกั้น) สามารถแบ่งออกเป็นข้อมูลที่ต่อเนื่อง (continuous data) และข้อมูลเป็นชิ้น/ไม่ต่อเนื่อง (discrete data) โดยวิธีการทางคณิตศาสตร์ที่ใช้วิเคราะห์ข้อมูลเหล่านี้ต่างกัน



Binomial Distribution Discrete Data & Discrete Probability Curve

Standard Normal Distribution Continuous Data and Continuous Probability Curve

www.statisticsfromatoz.com

ตัวอย่างฟังก์ชันทางคณิตศาสตร์ที่ใช้ในการวิเคราะห์ข้อมูลเหล่านี้ (ตอนนี้อาจจะยังอ่านไม่เข้าใจ แต่ภายหลังกลับมาทำความเข้าใจใหม่ได้) มีดังนี้

วิธีการ	Discrete Data	Continuous Data
ผลรวม/พื้นที่รวม	$S = \sum_{i=1}^{N} f(x_i) \cdot \Delta x$	$S = \int_0^N f(x)\ \mathrm{d}x$
ค่าเฉลี่ย	$\bar{x} = \frac{\sum_{i=1}^{N} f(x_i) \cdot \Delta x}{\sum_{i=1}^{N} \Delta x}$	$\bar{x} = \frac{1}{N} \int_0^N f(x)\ \mathrm{d}x$
ค่า RMS	$x_{\mathrm{RMS}} = \sqrt{\frac{1}{N} \sum_{i=1}^{N} x_i^2}$	$x_{\mathrm{RMS}} = \sqrt{\frac{1}{N} \int_0^N [f(x)]^2\ \mathrm{d}x}$
ค่าเฉลี่ยถ่วงน้ำหนัก	$\bar{x} = \frac{\sum_{i=1}^{N} w_i f(x_i) \cdot \Delta x}{\sum_{i=1}^{N} w_i \Delta x}$	N/A

การแจกแจงความถี่ (Frequency Distribution)

ตารางแจกแจงความถี่ (Frequency (distribution) table) คือ ตารางที่บ่งบอกความถี่ของข้อมูลในช่วงนั้น
ความถี่ (frequency) คือ จำนวนการเกิดซ้ำ ๆ หรือจำนวนที่พบ (number of occurrence)

ช่วงคะแนน	ความถี่ (f)	ความถี่สะสม (F)	ความถี่สัมพัทธ์	ความถี่สะสมสัมพัทธ์
1–10	3			
11–20	12			
21–30	15			
31–40	24			
41–50	6			
$N =$	60	sum =		

ตารางแสดงข้อมูลช่วงคะแนนของนักเรียนกลุ่มหนึ่ง

<u>การหาค่าต่าง ๆ</u>

1. ความถี่สะสม $F_i = F_{i-1} + f_i$
2. ความถี่สัมพัทธ์ $f_{\text{rel}} = \frac{f}{N}$

<u>ศัพท์ในตารางแจกแจงความถี่ (Technical formula), ใช้จริงไม่ต้องจำ (Sense)</u>

1. ขอบล่าง (Lower boundary)

$$L = \frac{\min_i + \max_{i-1}}{2}$$

2. ขอบบน (Upper boundary)

$$U = \frac{\max_i + \min_{i+1}}{2}$$

3. จุดกึ่งกลางของอันตรภาคชั้น (Midpoint/Midrange)

$$a = \frac{\min_i + \max_i}{2} = \frac{L + U}{2}$$

4. ความกว้างของอันตรภาคชั้น (Interval)

$$I = U - L = \max_i - \max_{i-1} = \min_i - \min_{i-1}$$

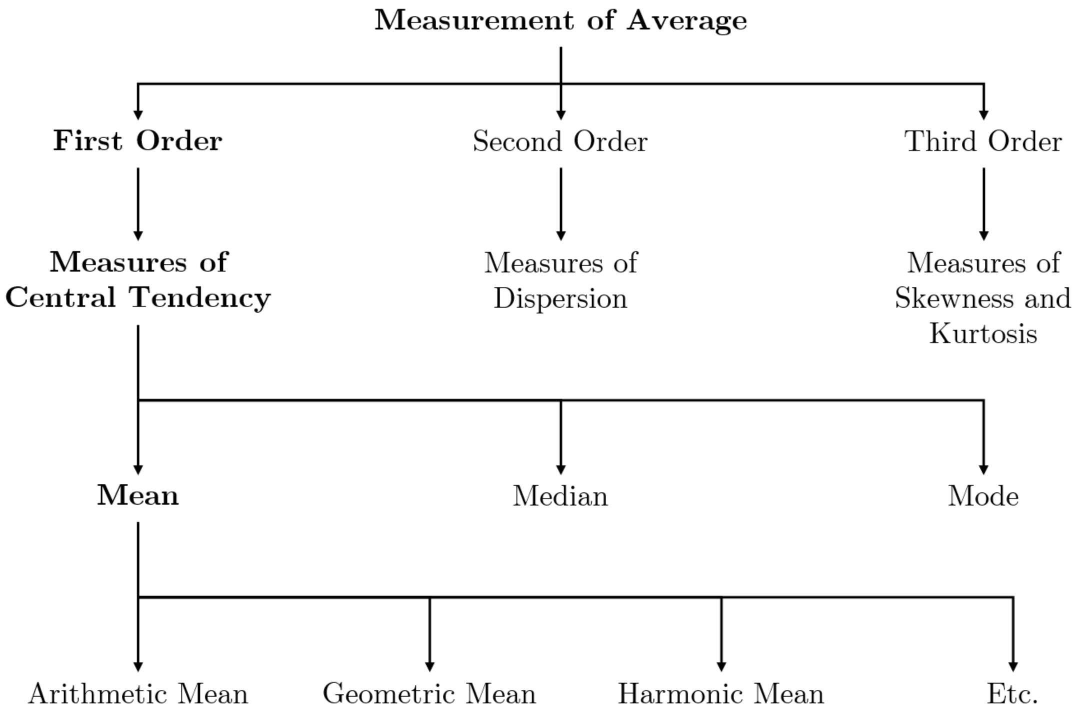
ช่วงคะแนน	ความถี่ (f)
1–10	3
11–20	12
21–30	15
31–40	24
41–50	6

การวาดกราฟแจกแจงความถี่ (Histogram-->Frequency polygon-->Frequency curve)

การวาดกราฟความถี่สะสม (Ogive)

ค่ากลางของข้อมูล (Intermediate value/Central tendency)

ค่ากลางของข้อมูล คือ ค่าที่สามารถใช้เป็นตัวแทนของกลุ่มข้อมูลนั้น ๆ ได้ ซึ่งมีประโยชน์ต่อการคำนวณ การประเมินข้อมูล การตัดสินใจต่าง ๆ เป็นภาพใหญ่ โดยไม่คำนึงถึงรายละเอียดข้อมูลรายตัว



ค่าเฉลี่ย (Mean) คือ ค่าที่อยู่ตรงกลาง โดยใช้ ค่าของข้อมูล เป็นตัววัด
มัธยฐาน (Median) คือ ค่าที่อยู่ตรงกลาง โดยใช้ ตำแหน่งของข้อมูล เป็นตัววัด
ฐานนิยม (Mode) คือ ค่าที่พบซ้ำมากที่สุดของข้อมูลนั้น

ค่ากลางแต่ละชนิดสามารถนำไปใช้ประโยชน์ได้แตกต่างกัน ขึ้นอยู่กับความเหมาะสมของการใช้งานกับลักษณะของข้อมูลนั้น ๆ ซึ่งบางข้อมูล ก็อาจหาค่ากลางบางชนิดไม่ได้ หรือแม้กระทั่งมีค่ากลางเท่ากัน

<u>การใช้สัญลักษณ์ทางคณิตศาสตร์เบื้องต้น</u>

สัญลักษณ์ทางคณิตศาสตร์หลาย ๆ สัญลักษณ์ ซึ่งถูกนำมาใช้เป็นทั้งเครื่องหมายการดำเนินการ (operator) และสัญลักษณ์ของตัวแปรหรือตัวถูกดำเนินการ (variable/operand) เป็นตัวอักขระกรีก (Greek symbol) อักขระคล้ายกรีก อักขระกรีกแบบตัวเขียน และ อักขระพิมพ์เขียน (Scripts Letter) เนื่องจากบางครั้งตัวแปรที่เขียนในสัญลักษณ์ละติน (Latin symbol) เช่น A-Z, 0-9 ไม่พอต่อความต้องการ เพราะความหมายอาจถูกมองพลาดและสับสนในตัวแปรไปมากหากใช้สัญลักษณ์เดียวกัน เช่น อุณหภูมิในระบบหน่วยวัดระหว่างประเทศ (The International System of Units, SI Unit) และระบบหน่วย CGS เป็นหน่วยเคลวิน (K) ใช้สัญลักษณ์แทนค่านั้นว่า T ส่วนอุณหภูมิในระบบหน่วยวัดเมตริก (Metric System, MKS) เป็นหน่วยองศาเซลเซียส (°C) ใช้สัญลักษณ์ t แต่ในเมื่อ t ก็มีความหมายว่าเวลา คนส่วนใหญ่จึงปรับอุณหภูมิใน MKS ให้มีสัญลักษณ์กรีกแบบตัวเขียน Theta ϑ แทน แต่ถึงกระนั้น ตัวแปร T ก็หมายถึงแรงตึงเชือก และคาบการกลับไปกลับมาได้เหมือนกัน

รวมถึงวิธีการเขียนสัญลักษณ์ทางคณิตศาสตร์อย่างถูกต้อง ทั้งเรื่องการห้อย การชี้ ตัวตรง ตัวเอียง ตัวหนา เมื่อเขียนแตกต่างกันย่อมมีความหมายแตกต่างกัน เช่น
- **ตัวหนา** ใช้บอกความเป็นเวกเตอร์และเมทริกซ์ ($\mathbf{A}$)
- *ตัวเอียง* ใช้เขียนสัญลักษณ์ตัวแปรทั่วไปที่บ่งบอกความเป็นปริมาณ (a, x, z, i)
- ตัวตรง ใช้บอกอักษรย่อ ตำแหน่ง สัญลักษณ์ดิฟเฟอร์เรนเชียล ชื่อฟังก์ชัน/ความสัมพันธ์ในบางกรณี คำเต็มและตัวเลข ($1, 2, \mathrm{RMS}, U_{\mathrm{A}}, \sin \omega t$)

แนะนำให้ผู้อ่านศึกษาข้อมูลเพิ่มในเรื่องการวัด ระบบหน่วยวัด ความเป็นมาของอักขระ
วงการคณิตศาสตร์และพิสิกส์ ซึ่งทั่วไปมักไม่มีใครมาเล่นให้ฟังโต้ง ๆ อยู่แล้ว แบบว่า อยู่ ๆ มีคนมาทักว่า
นายใช้ CGS หรือ MKS อะ ถ้ามีคนมาถามหน้าเซเว่นก็แปลกชอบกลอยู่

1. เครื่องหมาย Sigma (Σ) <--"S พิมพ์ใหญ่ ถ้าพิมพ์เล็ก s จะเป็น σ"

เครื่องหมาย Σ ใช้เป็นสัญลักษณ์แทนความหมายว่าการบวกทางพีชคณิต (arithmetic summation) ซึ่งมีสมบัติที่บังเอิญถูกค้นพบ/คิดค้นขึ้นมาต่าง ๆ (หาอ่านหาเรียนได้ทั่วไป) โดยมีสัญลักษณ์การใช้งานทั่วไป (generalized form) คือ

$$S = \sum_{i=m}^{N} (k^a b^c - 1) x_i y^i$$

2. เครื่องหมาย Pi (Π) ไม่ใช่ pi (π)

เครื่องหมาย ใช้เป็นสัญลักษณ์แทนการคูณทางพิชคณิต (arithmetic multiplication) เช่น

$$P = \prod_{i=m}^{N} (k^{abc} e^{x_i} + x_i)$$

<u>ค่าเฉลี่ย</u>

ค่าเฉลี่ย เกิดขึ้นจากการนำค่ามา “รวม” แล้ว “เฉลี่ย” ด้วยจำนวนตัว ซึ่งวิธีการ “รวม” นั้นมีแตกต่างวิธีกันไป ซึ่งมักสอดคล้องกันกับวิธีการ “เฉลี่ย” ที่แตกต่างกันตามมาด้วย โดยค่าเฉลี่ยที่มักพบบ่อยในการเรียนเนื้อหาทั่วไป จะมี 3 แบบ** (วิธีที่จะกล่าวเป็นการหาค่าเฉลี่ยของข้อมูลไม่ต่อเนื่อง):

1. ค่าเฉลี่ยเลขคณิต (Arithmetic Mean: AM)

$$\bar{x} = \mu = \mathrm{AM} = \frac{1}{n}\sum_{i=1}^{n} x_i = \frac{x_1 + x_2 + x_3 + \cdots + x_n}{n}$$

2. ค่าเฉลี่ยเรขาคณิต (Geometric Mean: GM)

$$\mathrm{GM} = \sqrt[n]{\prod_{i=1}^{n} x_i} = \sqrt[n]{x_1 \cdot x_2 \cdot x_3 \cdots x_n}$$

3. ค่าเฉลี่ยฮาร์โมนิค (Harmonic Mean: HM)

$$\frac{n}{\mathrm{HM}} = \sum_{i=1}^{n} \frac{1}{x_i} = \frac{1}{x_1} + \frac{1}{x_2} + \frac{1}{x_3} + \cdots + \frac{1}{x_n}$$

ในการประยุกต์ใช้นั้น ค่าเฉลี่ยรูปแบบอื่น ๆ จะถูกนำมาใช้ตามความเหมาะสม จึงยกตัวอย่างค่าเฉลี่ยต่าง ๆ ให้ผู้อ่านศึกษาเพิ่มเติมในเรื่องที่มา การใช้งาน และการประยุกต์ ดังนี้

4. ค่าเฉลี่ยเลขคณิตถ่วงน้ำหนัก (Weighted Arithmetic Mean)

$$\bar{x} = \frac{\sum_{i=1}^{n} w_i x_i}{\sum_{i=1}^{n} w_i} = \frac{w_1 x_1 + w_2 x_2 + w_3 x_3 + \cdots + w_n x_n}{w_1 + w_2 + w_3 + \cdots + w_n}$$

5. ค่าเฉลี่ยเรขาคณิตถ่วงน้ำหนัก (Weighted Geometric Mean)

$$\bar{x} = \left(\prod_{i=1}^{n} x_i^{w_i}\right)^{\frac{1}{\sum_{i=1}^{n} w_i}} = \sqrt[(w_1+w_2+w_3+\cdots+w_n)]{x_1^{w_1} \cdot x_2^{w_2} \cdot x_3^{w_3} \cdots x_n^{w_n}}$$

6. Quadratic Mean (Root-Mean-Square/RMS)

$$x_{\mathrm{RMS}} = \sqrt{\frac{1}{n}\sum_{i=1}^{n} x_i^2} = \sqrt{\frac{x_1^2 + x_2^2 + x_3^2 + \cdots + x_n^2}{n}}$$

7. Cubic Mean

$$\bar{x}_{\mathrm{cubic}} = \sqrt[3]{\frac{1}{n}\sum_{i=1}^{n} x_i^3} = \sqrt[3]{\frac{x_1^3 + x_2^3 + x_3^3 + \cdots + x_n^3}{n}}$$

8. Generalized Mean (Power Mean/Hölder mean)

$$M_p(S) = \sqrt[p]{\frac{1}{n}\sum_{i=1}^{n} x_i^p} = \left(\frac{1}{n}\sum_{i=1}^{n} x_i^p\right)^{\frac{1}{p}}; x_i \in S$$

9. Weighted Power Mean

$$M_p(S) = \sqrt[p]{\frac{1}{n}\sum_{i=1}^{n} w_i x_i^p} = \left(\frac{1}{n}\sum_{i=1}^{n} w_i x_i^p\right)^{\frac{1}{p}}; x_i \in S$$

For 8. And 9., there are special cases which has its name and usage. The generalized terms can be applied to any cases but with more difficult approach:

Given $S_N = \{x_1, x_2, x_3, \dots, x_n\}$.

n	Term/Definition*	Name
$-\infty$	$M_{-\infty}(S_n) = \lim_{p\to-\infty} M_p = \min(S_n)$	Minimum
-1	$M_{-1}(S_n) = \frac{n}{\sum_{i=1}^{n}\frac{1}{x_i}}$	Harmonic Mean
0	$M_0(S_n) = \lim_{p\to 0} M_p = \sqrt[n]{\prod_{i=1}^{n} x_i}$	Geometric Mean
1	$M_1(S_n) = \frac{1}{n}\sum_{i=1}^{n} x_i$	Arithmetic Mean
2	$M_2(S_n) = \sqrt{\frac{1}{n}\sum_{i=1}^{n} x_i^2}$	Root-Mean-Square
3	$M_3(S_n) = \sqrt[3]{\frac{1}{n}\sum_{i=1}^{n} x_i^3}$	Cubic Mean
$+\infty$	$M_{-\infty}(S_n) = \lim_{p\to-\infty} M_p = \max(S_n)$	Maximum

* For the proofs of equality/derivation from definition, they require higher level of mathematics (Calculus 2 and up) to fully understand. Studying calculus prior to entering university level is **highly** recommended. You can apply knowledge very early in scientific and engineering field of study.

10. Interquartile Mean

$$x_{\mathrm{IQM}} = \frac{2}{n}\sum_{i=\frac{n}{4}+1}^{\frac{3n}{4}} x_i$$

11. Midrange

$$M = \frac{\max x + \min x}{2}$$

12. Quasi-Arithmetic Mean (Generalized *f*-mean)

$$M_f(\vec{x}) = f^{-1}\left(\frac{1}{n}\sum_{k=1}^{n} f(x_k)\right) ; f \text{ is injective (one to one)}.$$

13. Trimean

$$\mathrm{TM} = \frac{1}{2}\left(Q_2 + \frac{Q_1 + Q_3}{2}\right)$$

14. Normalized Mean (This one is a large type consisting of many subtypes)

** โดยความสัมพันธ์ของค่าเฉลี่ยประเภท 1.–3. นั้นมีความสัมพันธ์ที่เป็นความจริงเสมอ (การพิสูจน์ใช้สมบัติทาง Sigma และ Pi และลอการิทึม หรือวิธีอื่น ๆ ตามสมควร) นำมาใช้ประยุกต์ในทฤษฎีทางพีชคณิตและทางเรขาคณิตได้มากมาย (*ศึกษาเพิ่มเติม ส่วนใหญ่เอามาใช้แก้โจทย์เลข ม.ต้น*) นั้นคือ

$$\mathrm{AM} \geq \mathrm{GM} \geq \mathrm{HM}$$

ในส่วนการคิด AM ธรรมดานั้น อาจมองอีกรูปเป็นอีกความหมายได้เป็น $\Sigma x = n \cdot \bar{x}$

<u>ค่าเฉลี่ย (อันตรภาคชั้นเป็นช่วง)</u>

ใช้จุดกึ่งกลางชั้นแล้วคิดแบบ Weighted Arithmetic Mean ได้เลย นั่นคือ

$$\bar{x} = \frac{\sum_{i=1}^{k} f_i x_i}{n} \;; n = \sum_{i=1}^{k} f_i = F_k$$

สำหรับสูตรเพื่อใช้ลดทอนข้อมูลแบบเป็นตารางแจกแจงความถี่แบบเป็นช่วง เวลาใช้สอบแล้วต้องคำนวณด้วยมือ แล้วเลขมันเยอะ สามารถใช้สูตรนี้ได้ (d คือชั้นสมมติ เป็นเลขอะไรก็ได้เรียงกัน):

$$\bar{x} = a + I\bar{d} \;; \bar{d} = \frac{\sum_{i=1}^{k} f_i d_i}{n}$$

<u>สมบัติของค่าเฉลี่ย</u>

1. $\Sigma(x_i - \bar{x}) = 0$
2. $\Sigma(x_i - \bar{x})^2 = \min \Sigma(x_i - a)^2 \;\; ; a \in \mathbb{R}$

มัธยฐาน

มัธยฐาน คือ ค่าที่เกิดจากการนำข้อมูลมาเรียงแถวกันตามลำดับค่าของมัน แล้วตัวที่อยู่ตำแหน่งตรงกลางเป๊ะ (dead center) จะเป็นมัธยฐานของข้อมูล ไม่ว่าข้อมูลจะมีจำนวนข้อมูลเป็นเลขคู่หรือเลขคี่ก็ตาม

$$2, 8, 18, 32, 50$$

$$1, 3, 5, 6, 7, 8, 12, 16, 20, 21, 27, 32, 33, 33, 34, 35$$

การหาตำแหน่งมัธยฐานของข้อมูล n จำนวน

$$\text{medpos} = \frac{n+1}{2}$$

มัธยฐาน (อันตรภาคชั้นเป็นช่วง)

ใช้สูตร (ซึ่งมีที่มาจาก sense แต่ รร.ไม่สอน)

$$\text{medpos} = \frac{n}{2}$$

$$\text{median} = L + \frac{I}{f}\left(\frac{n}{2} - \Sigma f_L\right) = U - \frac{I}{f}\left(\Sigma f - \frac{n}{2}\right)$$

สมบัติของมัธยฐาน

1. $\Sigma|x_i - \text{median}| = \min \Sigma|x_i - b| \;\; ; b \in \mathbb{R}$

ฐานนิยม

เอาตามตรง หัวข้อนี้ไม่มีอะไรต้องพูด กลับไปดูความหมาย เพราะในเรื่อง ฐานนิยม ยังมีข้อถกเถียงอยู่มากมาย ว่ามีเกินกี่ค่านับว่าไม่มีฐานนิยม (ซึ่งต่างจากหาค่าฐานนิยมไม่ได้เพราะทุกตัวเหมือนกัน) แต่ละตำราเขียนมาแตกต่างกัน แต่เอาเข้าจริงตอนเราทำข้อมูลเราไม่ได้โฟกัสค่ากลางชนิดนี้มากเท่าไรนัก

ฐานนิยม (อันตรภาคชั้นเป็นช่วง)

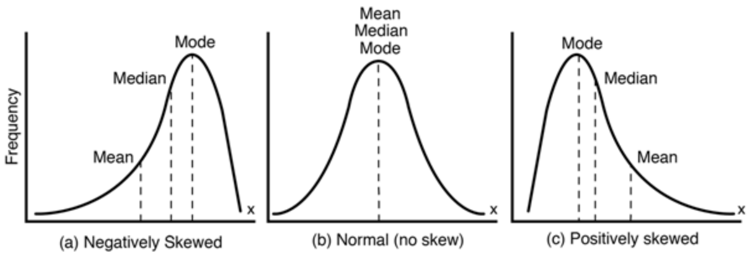
$$\text{mode} = L + I\left(\frac{d_L}{d_L + d_U}\right) = U - I\left(\frac{d_U}{d_L + d_U}\right)$$

<u>ความสัมพันธ์ระหว่าง ค่าเฉลี่ย-มัธยฐาน-ฐานนิยม</u>

มีคนเคยไว้ว่า ทั้ง 3 ค่ามีความสัมพันธ์แบบคร่าว ๆ และเป็นเชิงประมาณเหตุการณ์ (Empirical formula) ดังนี้

$$\bar{x} - \mathrm{mode} = 3(\bar{x} - \mathrm{median})$$

ซึ่งลักษณะการเรียงลำดับของทั้ง 3 ตัวสามารถบอกลักษณะของความเบ้ (Skew) ได้ (แต่บอกปริมาณของการเบ้ (Skewness & Kurtosis) ซึ่งเป็นการวัดความเฉลี่ยลำดับที่ 3 ไม่ได้ทุกกรณี)



Marco T. C. Faria: (a) เบ้ซ้าย (b) สมมาตร (c) เบ้ขวา

การวัดตำแหน่งของข้อมูล

ในหัวข้อก่อนหน้า มัธยฐาน (Median) เป็นหนึ่งในการวัดตำแหน่งของข้อมูล ซึ่งอยู่ตำแหน่งตรงกลาง แต่เราสามารถวัดหาตำแหน่งใด ๆ เช่น ที่ 70% ของข้อมูลมีค่าเท่าไร หรือ ที่สิ้นไตรมาสแรกข้อมูลมีค่าเท่าไร ฯลฯ

มัธยฐาน (Median)

median

$$10\ 11\ 15\ 17\ 18\ 21\ 26\ 27\ 32\ \underline{\mathbf{33}}\ 35\ 39\ 40\ 43\ 44\ 45\ 46\ 49\ 51$$

ควอร์ไทล์ (Quartile), เดไซล์ (Decile) และ เปอร์เซ็นไทล์ (Percentile)

Q_0 Q_1 Q_2 Q_3 Q_4

$$10\ 11\ 15\ 17\ \underline{\mathbf{18}}\ 21\ 26\ 27\ 32\ \underline{\mathbf{33}}\ 35\ 39\ 40\ 43\ \underline{\mathbf{44}}\ 45\ 46\ 49\ 51$$

$$10\ 11\ 15\ 17\ 18\ 21\ 26\ 27\ 32\ 33\ 35\ 39\ 40\ 43\ 44\ 45\ 46\ 49\ 51$$

$$10\ 11\ 15\ 17\ 18\ 21\ 26\ 27\ 32\ 33\ 35\ 39\ 40\ 43\ 44\ 45\ 46\ 49\ 51$$

การหาตำแหน่งของข้อมูลที่เรียงเดี่ยว (ไม่เป็นช่วง) สามารถหาได้จาก

$$p(Q_r) = \frac{r}{4}(n+1)$$

$$p(D_r) = \frac{r}{10}(n+1)$$

$$p(P_r) = \frac{r}{100}(n+1)$$

สำหรับความสัมพันธ์แบบ n-tile จะได้ว่า

ในการหา n-tile เราสามารถเริ่มจากการหา “ตำแหน่ง” แล้วจึงหา “ค่าที่ตำแหน่งนั้นได้”

กรณี n-tile ไม่ใช่จำนวนเต็ม ให้นำส่วนทศนิยม (Fractional part) มาคูณกับผลต่างของสองตำแหน่งที่ n-tile อยู่ระหว่าง ซึ่งคือ ...

<u>ตัวอย่าง</u> 20 22 25 25 27

29 31 33 33 34

38 39 39 41 45

48 49 51 54 55

57 59 59

<u>ควอร์ไทล์ (Quartile), เดไซล์ (Decile) และ เปอร์เซ็นไทล์ (Percentile) (อันตรภาคชั้นเป็นช่วง)</u>

สำหรับ Q, D, P สำหรับอันตรภาคชั้นเป็นช่วง สามารถหาตำแหน่งได้จาก

$$p(Q_r) = \frac{r}{4}(n)$$

$$p(D_r) = \frac{r}{10}(n)$$

$$p(P_r) = \frac{r}{100}(n)$$

แล้วจึงนำมาหาค่าที่ตำแหน่งนั้นในสูตรที่ รร.สอนมา คล้าย ๆ กับ Median เพียงเปลี่ยนจากตำแหน่งกึ่งกลาง ($\frac{n}{2}$) เป็นตำแหน่งใด ๆ ที่เราต้องการหา

$$X_r = L + \frac{I}{f}(p(X_r) - \Sigma f_L) = U - \frac{I}{f}\big(\Sigma f - p(X_r)\big)$$

References

[1] RATH Center. Statistics. https://rathcenter.com/old-web/Sheet/Sttstc.pdf

[2] OpenStax. Introductory Statistics. https://d3bxy9euw4e147.cloudfront.net/oscms-prodcms/media/documents/IntroductoryStatistics-LR.pdf

[3] University of Toronto. Probabilities and Statistics (2ed). http://www.utstat.toronto.edu/mikevans/jeffrosenthal/book.pdf

[4] Economics Discussion. Central Tendency. https://www.economicsdiscussion.net/central-tendency/central-tendency-a-close-view/12182